

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



AF

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 9/90, 15/61, A61K 37/58	A1	(11) International Publication Number: WO 95/31538 (43) International Publication Date: 23 November 1995 (23.11.95)
(21) International Application Number: PCT/US94/05701 (22) International Filing Date: 18 May 1994 (18.05.94) (71) Applicant (for all designated States except US): HUMAN GENOME SCIENCES, INC. [US/US]; Suite 300, 9620 Medical Center Drive, Rockville, MD 20850-3358 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): WEI, Ying-Fei [CN/US]; 13524 Straw Bale Lane, Darnestown, MD 20878 (US). ADAMS, Mark, D. [US/US]; 15205 Duffief Drive, North Potomac, MD 20878 (US). FLEISCHMANN, Robert, D. [US/US]; 2800 Woodley Road, NW #134, Washington, DC 20008 (US). (74) Agents: OLSTEIN, Elliot, M. et al.; Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein, 6 Becker Farm Road, Roseland, NJ 07069-1739 (US).		(81) Designated States: AU, CA, CN, JP, KR, NZ, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: HUMAN DNA TOPOISOMERASE I-ALPHA (57) Abstract Disclosed is a human hTopI- α polypeptide and DNA (RNA) encoding such hTopI- α polypeptide. Also provided is a procedure for producing such polypeptide by recombinant techniques and antibodies and antagonists/inhibitors against such polypeptide. Also provided are methods of using the antibodies and antagonist inhibitors to inhibit the action of hTopI- α for therapeutic purposes such as an antitumor agent, to detect an autoimmune disease, or retroviral infections and to treat adenocarcinoma of the colon.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LU	Luxembourg	SN	Senegal
CN	China	LV	Latvia	TD	Chad
CS	Czechoslovakia	MC	Monaco	TG	Togo
CZ	Czech Republic	MD	Republic of Moldova	TJ	Tajikistan
DE	Germany	MG	Madagascar	TT	Trinidad and Tobago
DK	Denmark	ML	Mali	UA	Ukraine
ES	Spain	MN	Mongolia	US	United States of America
FI	Finland			UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

HUMAN DNA TOPOISOMERASE I - ALPHA

This invention relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides, the use of such polynucleotides and polypeptides, as well as the production of such polynucleotides and polypeptides. More particularly, the polypeptide of the present invention is human DNA topoisomerase I alpha (hTopI- α). The invention also relates to inhibiting the action of such polypeptides.

DNA topoisomerase I and II catalyze the breaking and rejoining of DNA strands in a way that allows the strands to pass through one another, thus altering the topology of DNA. Type I topoisomerase recognizes double-stranded DNA but only breaks on strand in the process of relaxing DNA, while the type II enzyme breaks both strands of duplex DNA. Both enzymes can perform a variety of similar topological inter-conversions, including relaxation of super coiled DNA, knotting and unknotting and catenation and decatenation of duplex DNA. Topoisomerase I is ATP-independent, while Topoisomerase II requires energy.

Both topoisomerase I and II can provide the topological inter-conversions necessary for transcription and replication. For example, topoisomerase I can provide the necessary unlinking activity for efficient *in vitro* DNA

replication (Minden, et al., J. Biol. Chem., 260:9316, (1985)), however, topoisomerase II can also facilitate the replication of SV40 DNA by HeLa cell lysates (Yang, et al., Proceedings of the National Academy of Sciences, U.S.A., 84:950, (1987)). Genetic studies in yeast reveal that both replication and transcription proceed in single mutants deficient in either topoisomerase I or II (Goto, et al., Proceedings of the National Academy of Sciences, U.S.A., 82:7178 (1985)). In cells lacking both topoisomerases, transcription and replication are dramatically reduced (Uemura, et al., EMBO Journal, 5:1003 (1986)).

Several lines of evidence suggest that topoisomerase I normally functions during transcription. The enzyme has been shown to be localized preferentially to actively transcribed loci by immunofluorescence (Fleishmann, et al., Proceedings of the National Academy of Sciences, U.S.A., 81:6958 (1984)), and by co-immunoprecipitation with transcribed DNA (Gilmore, et al., Cell, 44:401, (1986)). Furthermore, topoisomerase I cleavage sites have been mapped to regions in and around transcribed DNA (Bonner, et al., Cell, 41:541 (1985)). Nonetheless, at least in yeast, topoisomerase II can apparently substitute for the functions of topoisomerase I in transcription.

While all cells utilize Topoisomerase I and II for transcription and replication, cells with a high amount of transcription and replication, eg. cancerous cells, have a much higher concentration of Topoisomerase I and II.

Topoisomerase I has been used to classify autoimmune disease. Autoimmune diseases are diseases in which an animal's immune system attacks its own tissues. Often the various types of autoimmune disease can be characterized based upon the specificity of autoantibodies which are produced. For example, it is well known that the serum of patients having the connective tissue autoimmune disease progressive systemic sclerosis, also known as scleroderma,

frequently contain antibodies to such nuclear antigens--as topoisomerase I. Thus, the ability to accurately detect the presence of antibodies reactive with topoisomerase I can greatly assist in evaluating the prognosis and planning, or monitoring, of the appropriate therapy for patients with scleroderma.

A 3645-base pair human topoisomerase I cDNA clone and a mutated version of the cDNA encoding a protein with phenylalanine instead of tyrosine at position 723 have been overexpressed two to five fold in stably transfected baby hamster kidney cells. The results of this overexpression indicate that tyrosine 723 is essential for enzyme activity and is consistent with predictions based on homology comparisons with the yeast enzymes, that this is the active-site tyrosine in the human topoisomerase I. (Madden, K.R. and Champoux, J.J., Cancer Research, 52:525-532, (1992)).

Also, cDNA clones encoding human topoisomerase I have been isolated from an expression vector library screened with autoimmune anti-topoisomerase I serum. The sequence data shows that the catalytically active 67.7-kDa fragment is comprised of the carboxyl terminus, (D'Arpa, P. et al., Proc. Natl. Acad. Sci. U.S.A., 85:2543-2547, (1988)).

cDNA molecules coding for eukaryotic topoisomerase I polypeptide which encode at least one epitope for autoantibodies to eukaryotic topoisomerase I and cloning vehicles capable of expressing these cDNA molecules are disclosed in United States Patent No. 5,070,192.

In accordance with one aspect of the present invention, there are provided novel mature polypeptides which are hTopI- α , as well as fragments, analogs and derivatives thereof. The polypeptides of the present invention are of human origin.

In accordance with another aspect of the present invention, there are provided polynucleotides (DNA or RNA) which encode such polypeptides.

In accordance with yet a further aspect of the present invention, there is provided a process for producing such polypeptides by recombinant techniques.

In accordance with yet a further aspect of the present invention, there are provided antibodies against such polypeptides, which may be used diagnostically to detect cancers and autoimmune diseases.

In accordance with yet another aspect of the present invention, there are provided antagonist/inhibitors to such polypeptides, which may be used therapeutically to inhibit the action of such polypeptides, for example, as an antitumor agent and to detect autoimmune diseases, and to treat adenocarcinoma of the colon, and retroviral infections.

These and other aspects of the present invention should be apparent to those skilled in the art from the teachings herein.

The following drawings are illustrative of embodiments of the invention and are not meant to limit the scope of the invention as encompassed by the claims.

Fig. 1 shows the cDNA and corresponding deduced amino acid sequence of hTopI- α . The polypeptide encoded by the amino acid sequence shown is the mature form of the polypeptide. The standard one letter abbreviation for amino acids is used.

Fig. 2 shows a comparison of hTOPI- α and human topoisomerase I at the amino acid level. The upper line is hTOPI- α .

In accordance with an aspect of the present invention, there is provided an isolated nucleic acid (polynucleotide) which encodes for the mature polypeptide having the deduced amino acid sequence of Figure 1 or for the mature polypeptide encoded by the cDNA of the clone deposited as ATCC Deposit No. 75714 on March 18, 1994.

A polynucleotide encoding a polypeptide of the present invention was obtained from a fetal brain cDNA library. It

is homologous to human topoisomerase I. It contains an open reading frame encoding a protein of approximately 601 amino acid residues and it is structurally related to human DNA topoisomerase I showing 86% similarity and 70% identity at the amino acid level. Further, hTopI- α shows 83% similarity and 67% identity to human topoisomerase I as published by D'Arpa et al. The amino acid tyrosine⁷²³.

The polynucleotide of the present invention may be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand. The coding sequence which encodes the mature polypeptide may be identical to the coding sequence shown in Figure 1 or that of the deposited clone or may be a different coding sequence which coding sequence, as a result of the redundancy or degeneracy of the genetic code, encodes the same, mature polypeptide as the DNA of Figure 1 or the deposited cDNA.

The polynucleotide which encodes for the mature polypeptide of Figure 1 or for the mature polypeptide encoded by the deposited cDNA may include: only the coding sequence for the mature polypeptide; the coding sequence for the mature polypeptide and additional coding sequence such as a leader or secretory sequence or a proprotein sequence; the coding sequence for the mature polypeptide (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequence 5' and/or 3' of the coding sequence for the mature polypeptide.

Thus, the term "polynucleotide encoding a polypeptide" encompasses a polynucleotide which includes only coding sequence for the polypeptide as well as a polynucleotide which includes additional coding and/or non-coding sequence.

The present invention further relates to variants of the hereinabove described polynucleotides which encode for fragments, analogs and derivatives of the polypeptide having

the deduced amino acid sequence of Figure 1 or the polypeptide encoded by the cDNA of the deposited clone. The variant of the polynucleotide may be a naturally occurring allelic variant of the polynucleotide or a non-naturally occurring variant of the polynucleotide.

Thus, the present invention includes polynucleotides encoding the same mature polypeptide as shown in Figure 1 or the same mature polypeptide encoded by the cDNA of the deposited clone as well as variants of such polynucleotides which variants encode for a fragment, derivative or analog of the polypeptide of Figure 1 or the polypeptide encoded by the cDNA of the deposited clone. Such nucleotide variants include deletion variants, substitution variants and addition or insertion variants.

As hereinabove indicated, the polynucleotide may have a coding sequence which is a naturally occurring allelic variant of the coding sequence shown in Figure 1 or of the coding sequence of the deposited clone. As known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which does not substantially alter the function of the encoded polypeptide.

The polynucleotides of the present invention may also have the coding sequence fused in frame to a marker sequence which allows for purification of the polypeptide of the present invention. The marker sequence may be a hexahistidine tag supplied by a pQE-9 vector to provide for purification of the mature polypeptide fused to the marker in the case of a bacterial host, or, for example, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson, I., et al., Cell, 37:767 (1984)).

The present invention further relates to polynucleotides which hybridize to the hereinabove-described

sequences if there is at least 50% and preferably 70% identity between the sequences. The present invention particularly relates to polynucleotides which hybridize under stringent conditions to the hereinabove-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences. The polynucleotides which hybridize to the hereinabove described polynucleotides in a preferred embodiment encode polypeptides which retain substantially the same biological function or activity as the mature polypeptide encoded by the cDNA of Figure 1 or the deposited cDNA.

The deposit(s) referred to herein will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for purposes of Patent Procedure. These deposits are provided merely as convenience to those of skill in the art and are not an admission that a deposit is required under 35 U.S.C. §112. The sequence of the polynucleotides contained in the deposited materials, as well as the amino acid sequence of the polypeptides encoded thereby, are incorporated herein by reference and are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited materials, and no such license is hereby granted.

The present invention further relates to a hTopI- α polypeptide which has the deduced amino acid sequence of Figure 1 or which has the amino acid sequence encoded by the deposited cDNA, as well as fragments, analogs and derivatives of such polypeptide.

The terms "fragment," "derivative" and "analog" when referring to the polypeptide of Figure 1 or that encoded by the deposited cDNA, means a polypeptide which retains essentially the same biological function or activity as such

polypeptide. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide.

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a synthetic polypeptide, preferably a recombinant polypeptide.

The fragment, derivative or analog of the polypeptide of Figure 1 or that encoded by the deposited cDNA may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a

composition, and still be isolated in that such vector or composition is not part of its natural environment.

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the hTopI- α genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The polynucleotides of the present invention may be employed for producing polypeptides by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing a polypeptide. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such

procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the E. coli. lac or trp, the phage lambda P_L promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in E. coli.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as E. coli, Streptomyces, Salmonella typhimurium; fungal cells, such as yeast; insect cells such as Drosophila and Sf9; animal cells such as CHO, COS or Bowes melanoma; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a

forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pbs, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are PKK232-8 and PCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P_R, P_L and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation. (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples including the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), α -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences,

and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include E. coli, Bacillus subtilis, Salmonella typhimurium and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well known to those skilled in the art.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

The polypeptides can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The polypeptides of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques

from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. Polypeptides of the invention may also include an initial methionine amino acid residue.

The polypeptides of the present invention are useful for identifying other molecules which have similar biological activity. An example of a screen for this is isolating the coding region of the hTopI- α gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention are used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

This invention also provides a method of screening drugs to identify drugs which specifically interact with, and bind to, hTopI- α which comprises contacting a mammalian cell comprising a DNA molecule encoding hTOPI- α with a plurality of drugs, detecting those drugs which bind to the mammalian cell, and thereby identifying drugs which specifically interact with, and bind to, hTOPI- α . Various methods of detection may be employed. The drugs may be "labeled" by association with a detectable marker substance (e.g., radiolabel or a non-isotopic label such as biotin). Drug candidates are identified by choosing chemical compounds which bind with high affinity to the expressed hTOPI- α protein in transfected cells, using radioligand binding methods well known in the art.

The sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few

chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the cDNA is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a cDNA clones to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 500 or 600 bases; however, clones larger than 2,000 bp have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. FISH requires use of the clones from which the EST was derived, and the longer the better. For example, 2,000 bp is good, 4,000 is better, and more than 4,000 is probably not necessary to get good results

a reasonable percentage of the time. For a review of this technique, see Verma et al., Human Chromosomes: a Manual of Basic Techniques, Pergamon Press, New York (1988).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region associated with the disease could be one of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb).

The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the polypeptides corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptides itself. In this manner, even a sequence encoding only a fragment of the polypeptides can be used to generate antibodies binding the whole native

polypeptides. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, 1975, *Nature*, 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, *Immunology Today* 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole, et al., 1985, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention.

The present invention also relates to a diagnostic assay for detecting the concentration of hTOPI- α in a sample from a host. An example of such an assay is an ELISA assay which utilizes an antibody specific to an hTOP-I α antigen, preferably a monoclonal antibody, which is coupled to an indicator enzyme such as horseradish peroxidase to generate a very specific and sensitive assay system. After binding of the peroxidase-coupled antibody to the antigen, the peroxidase can be used to generate a colored product that is measurable and whose concentration is related to the amount of antigen in a sample. Because of the catalytic nature of the enzyme the system greatly amplifies the signal. A high level of oxalyl-CoA decarboxylase is indicative of cancer since some human colon carcinoma cells have increased levels of hTopI- α . They may also be indicative of autoimmune diseases, such as scleroderma, rheumatoid arthritis and AIDS related complex.

Other examples of immunoassays which can be used to detect the level of antibodies against the polypeptide of the present invention are competitive and non-competitive

immunoassays in either a direct or indirect format. Examples include the radioimmunoassay, the sandwich (immunometric) assay and the Western blot assay. Detection of antibodies which bind to the hTopI- α of the invention can be done utilizing immunoassays which run in either the forward, reverse or simultaneous modes, including immunohistochemical assays on physiological samples. Regardless of the type of immunoassay used, the concentration of hTopI- α utilized can be readily determined by one of ordinary skill in the art using routine experimentation.

The hTopI- α of the invention can be labeled and bound to many different carriers and used to detect the presence of antibody specifically reactive with the polypeptide. Examples of well-known carriers include glass, polystyrene, polyvinyl chloride, polypropylene, polyethylene, polycarbonate, dextran, nylon, amylases, natural and modified cellulose, polyacrylamides, agarose and magnetite. The nature of the carrier can be either soluble or insoluble.

The present invention is also directed to antagonist/inhibitor molecules of the polypeptides of the present invention which may be used to reduce or eliminate the function of the polypeptide.

An example of an antagonist is an antibody, or in some cases, an oligonucleotide which binds to the hTopI- α polypeptides. An example of an inhibitor is a small molecule which binds to and occupies the catalytic site of the polypeptide thereby making the catalytic site inaccessible to substrate such that normal biological activity is prevented. Examples of small molecules include but are not limited to small peptides or peptide-like molecules.

Another example of an inhibitor is an antisense construct which inhibits hTopI- α *in vivo* by the use of antisense technology. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on

binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes for the mature polypeptides of the present invention, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix -see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al, Science, 241:456 (1988); and Dervan et al., Science, 251: 1360 (1991)), thereby preventing transcription and the production of hTopI- α . The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the hTopI- α (antisense - Okano, J. Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)).

The antagonist/inhibitors may be used to treat tumors since specific inhibition of htopI- α will inhibit tumor cell growth by blocking tumor cell DNA replication. The antagonist/inhibitors may also be used to treat retroviral infections by inhibiting hTopI- α and therefore blocking initiation and replication of the virus. The antagonist inhibitors may also be used to treat adenocarcinoma of the colon, since metastases are prevented by blocking DNA transcription of the cancerous cells.

The present invention is also directed to an assay using hTopI- α to identify antagonist/inhibitors of hTopI- α and/or human Topoisomerase I. DNA, hTopI- α and a potential antagonist/inhibitor could be combined together under appropriate conditions for a length of time sufficient for hTopI- α to act on the single strand DNA. The DNA could then be analyzed, for example, by gel electrophoresis, to determine whether hTopI- α functioned properly and in this way it could be determined whether there was an effective antagonist/inhibitor.

The compounds, e.g., antagonist/inhibitor compounds, of the present invention, may be employed in combination with a suitable pharmaceutical carrier. Such compositions comprise a therapeutically effective amount of the antagonist/inhibitor, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The formulation should suit the mode of administration.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

The pharmaceutical compositions may be administered in an effective amount to effectively inhibit hTopI- α from facilitating DNA transcription and replication, in the case where this action leads to undesired conditions, eg. retroviral infections.

The present invention will be further described with reference to the following examples; however, it is to be understood that the present invention is not limited to such examples. All parts or amounts, unless otherwise specified, are by weight.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available,

publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 μg of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 μl of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μg of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37°C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

Size separation of the cleaved fragments is performed using 8 percent polyacrylamide gel described by Goeddel, D. et al., *Nucleic Acids Res.*, 8:4057 (1980).

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T., et al., *Id.*, p. 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units to T4 DNA ligase ("ligase") per 0.5 μg of approximately equimolar amounts of the DNA fragments to be ligated.

Unless otherwise stated, transformation was performed as described in the method of Graham, F. and Van der Eb, A., Virology, 52:456-457 (1973).

Example 1

Expression of Recombinant hTopI- α in COS cells

The expression of plasmid, pcDNATopI HA is derived from a vector pcDNAI/Amp (Invitrogen) containing: 1) SV40 origin of replication, 2) ampicillin resistance gene, 3) E.coli replication origin, 4) CMV promoter followed by a polylinker region, a SV40 intron and polyadenylation site. A DNA fragment encoding the entire hTopI- α precursor and a HA tag fused in frame to its 3' end was cloned into the polylinker region of the vector, therefore, the recombinant protein expression is directed under the CMV promoter. The HA tag correspond to an epitope derived from the influenza hemagglutinin protein as previously described (I. Wilson, H. Niman, R. Heighten, A. Cherenson, M. Connolly, and R. Lerner, 1984, Cell 37, 767). The infusion of HA tag to our target protein allows easy detection of the recombinant protein with an antibody that recognizes the HA epitope.

The plasmid construction strategy is described as follows:

The expression plasmid pcDNATopI- α ATCC # 75714, encoding for hTopI- α was constructed by PCR on the pBLTopI- α using two primers: the 5' primer 5' - CGGGATCCATGCGCGTGGTGCGG - 3' contains a Bam HI site followed by 15 nucleotides of HhTopI- α coding sequence starting from the initiation codon; the 3' sequence 5' CGCTCTAGATCAAGCGTAGTCTGGGACGTCGTATGGGTAGAATTCAAAGTCTTCTCC - 3' contains complementary sequences to an Xba I site, translation stop codon, HA tag and the last 18 nucleotides of the hTopI- α coding sequence (not including the stop codon). Therefore, the PCR product contains a Bam HI site, active hTopI- α coding sequence followed by HA tag fused in frame, a

translation termination stop codon next to the HA tag, and an Xba I site. The PCR amplified DNA fragment and the vector, pCDNA1/Amp, were digested with Bam HI and Xba I restriction enzyme and ligated. The ligation mixture was transformed into E. coli strain SURE (available from Stratagene Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA 92037) the transformed culture was plated on ampicillin media plates and resistant colonies were selected. Plasmid DNA was isolated from transformants and examined by restriction analysis for the presence of the correct fragment. For expression of the recombinant hTopI- α , COS cells were transfected with the expression vector by DEAE-DEXTRAN method. (J. Sambrook, E. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Laboratory Press, (1989)). The expression of the hTopI- α HA protein was detected by radiolabelling and immunoprecipitation method. (E. Harlow, D. Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, (1988)). Cells were labelled for 8 hours with ^{35}S -cysteine two days post transfection. Culture media were then collected and cells were lysed with detergent (RIPA buffer (150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50mM Tris, pH 7.5). (Wilson, I. et al., Id. 37:767 (1984)). Both cell lysate and culture media were precipitated with a HA specific monoclonal antibody. Proteins precipitated were analyzed on 15% SDS-PAGE gels.

Example 2

In vitro transcription and translation of hTopI- α

The *in vitro* transcription and translation of the hTopI- α was carried out using the TNT Coupled Reticulocyte Lysate System (Promega, Madison, WI). The plasmid vector used is pBLSK. The cDNA encoding for hTopI- α was cloned directionally EcoRI to XhoI with the EcoRI site defining the 5' end of the gene and the XhoI site defining the 3' end of the gene. The gene was inserted in the T3 direction. T3

defines a bacteriophage RNA polymerase which recognizes a specific promoter, and transcribes the DNA into a mRNA. One microgram of the pBLSKhTOPI α was incubated with 25 μ l of TNT rabbit reticulocyte lysate, 2 μ l TNT reaction buffer, 1 μ l T3 RNA polymerase, 1 μ l of amino acid mixture minus methionine (1mM), 4 μ l of 35 S-methionine (1,000Ci/ μ mol) at 10mCi/ml, 1 μ l RNasin ribonuclease inhibitor (40U/ μ l) in 50 μ l of final volume at 37°C for 1.5 hour. 5 μ l of the reaction mixture was mixed with loading buffer, boiled for 5 minutes and loaded on a 10% SDS polyacrylamide gel to separate the protein. The gel was then fixed 10% acetic acid, 10% methanol at room temperature for 30 minutes, soaked in Amplify solution (Amersham) at room temperature for 1.5 hours, dried, and subjected to autoradiograph. The observed molecular weight of the hTopI- α in this system is 70 kD, which agrees with the predicted molecular weight by the sequence.

Example 3

Expression pattern of hTopI- α in human tissue

Northern blot analysis was carried out to examine the levels of expression of hTopI α in human tissues. Total cellular RNA samples were isolated with RNazol™ B system (Biotech Laboratories, Inc. 6023 South Loop East, Houston, TX 77033). About 10 μ g of total RNA isolated from each human tissue specified was separated on 1% agarose gel and blotted onto a nylon filter. (Sambrook, Fritsch, and Maniatis, Molecular Cloning, Cold Spring Harbor Press, (1989)). The labeling reaction was done according to the Stratagene Prime-It kit with 50ng DNA fragment. The labeled DNA was purified with a Select-G-50 column. (5 Prime - 3 Prime, Inc. 5603 Arapahoe Road, Boulder, CO 80303). The filter was then hybridized with radioactive labeled full length hTopI- α gene at 1,000,000 cpm/ml in 0.5 M NaPO $_4$, pH 7.4 and 7% SDS overnight at 65°C. After wash twice at room temperature and

twice at 60°C with 0.5 x SSC, 0.1% SDS, the filter was then exposed at -70°C overnight with an intensifying screen. The message RNA for hTopI α is present in all the tissues with abundance in ovary, testes, lung, spleen and prostate.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, within the scope of the appended claims, the invention may be practiced otherwise than as particularly described.

SEQUENCE LISTING

- (1) GENERAL INFORMATION:
- (i) APPLICANT: WEI, ET AL.
- (ii) TITLE OF INVENTION: Human Topoisomerase I- α
- (iii) NUMBER OF SEQUENCES: 2
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
CECCHI, STEWART & OLSTEIN
- (B) STREET: 6 BECKER FARM ROAD
- (C) CITY: ROSELAND
- (D) STATE: NEW JERSEY
- (E) COUNTRY: USA
- (F) ZIP: 07068
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 3.5 INCH DISKETTE
- (B) COMPUTER: IBM PS/2
- (C) OPERATING SYSTEM: MS-DOS
- (D) SOFTWARE: WORD PERFECT 5.1
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE: SUBMITTED HEREWITH
- (C) CLASSIFICATION:
- (vii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: FERRARO, GREGORY D.
- (B) REGISTRATION NUMBER: 36,134
- (C) REFERENCE/DOCKET NUMBER: 325800-104

(viii) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 201-994-1700

(B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 1,917 BASE PAIRS

(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: SINGLE

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GCAGATGCCG GTGGTGCGGC TGCTGCGGCT CCGGGCGGCT CTGACGCTGC TCGGGGAGGT 60
CCCCCGCCGC CCGGCTCCC GGGGTGTCCC GGGCTCGCGC AGGACGCAGA AGGGCAGTGG 120
AGCCAGGTGG GAGAAGGAGA AGCACGAAGA CGGGGTGAAG TGGAGACAGC TGGAGCACAA 180
GGGCCCCGAC TTCGCACCCC CATACHAGCC CCTTCCCGAC GGAGTGCGTT TCTTCTATGA 240
AGGAAGGCCT GTGAGATTGA GCGTGCCAGC GGAGGAGGTC CCCACTTTTT ATGGGAGGAT 300
GTTAGATCAT GAATACACAA CAAAGGAGGT TTTCCGGAAG AACTTCTTCA ATGACTGGCG 360
AAAGGAAATG GCGGTGGAAG AGAGGGAAGT CATCAAGAGC CTGGACAAGT GTGACTTCAC 420
GGAGATCCAC AGATACTTTG TGGACAAGGC CGCAGCCCGG AAAGTCCTGA GCAGGGAGGA 480
GAAGCAGAAG CTAAAAGAAG AGGCAGAAAA ACTTCAGCAA GAGTTCGGCT ACTGTATTTT 540
AGATGGTCAC CAAGAAAAAA TAGGCAACTT CAAGATTGAG CCGCCTGGCT TGTTCCGTGG 600
CCGTGGCGAC CATCCCAAGA TGGGGATGCT GAAGAGAAGG ATCACGCCAG AGGATGTGCT 660
TATCAACTGC AGCAGGGACT CGAAGATCCC CGAGCCGCCG GCGGGGCACC AGTGGAAGGA 720
GGTGCGCTCC GATAACACCG TCAXGTGGCT GGACAGCTTG ACCGAGAGCG TTCAGAACTC 780
CATCAAGTAC ATCATGCTGA ACCCTTGCTC GAAGCTGAAG GGGGAGACAG CTTGGCAGAA 840
GTTTGAAACA GCTCGACGCC TCGGGGGATT TGTGGACGAG ATCCGCTCCC AGTACCGGGC 900
TGAAGTGAAG TCTCGGAAA TGAAGACGAG ACAGCGGGCG GTGGCCCTGT ATTTTCATCGA 960
TAAGCTGGCA CTGAGAGCAG GAAATGAGAA GGAGGACGGT GAGGCGGCCG ACACCGTGGG 1020
CTGCTGTTCC CTCCGCGTGG AGCAGCTCCA GCTGCACCCG GAGGCCGATG GTTGCCAACA 1080
CGTGGTGGA TTTGACTTCC TGGGGAAGGA CTGCATCCGC TACTACAACA GAGTGCCGGT 1140
GGAGAAGCCG GTGTACAAGA ACTTACAGCT CTTTATGGAG AACAAAGACC CCCGGGACGA 1200
CCTCTTCGAC AGGCTGACCA CGACCAGCCT GAACAAGCAC CTCCAGGAGC TGATGGACGG 1260
```

GCTGSCGGCC AAGGTGTTCC GGACCTACAA CGCCTCCATC ACTCTGCAGG AGCAGCTGCG 1320
 GGCCCTGACG CGCGCCGAGG ACAGCATAGC AGCTAAGATC TTATCCTACA ACCGAGCCAA 1380
 CCGAGTCGTG GCCATTCTCT GCAACCATCA GCGAGCAACC CCCAGTACGT TCGAGAAGTC 1440
 GATGCAGAAT CTCCAGACGA AGATCCAGGC AAAGAAGGAG CAGGTGGCTG AGGCCAGGGC 1500
 AGAGCTGAGG AGGGCGAGGG CTGAGCACAA AGCCCAAGGG GATGGCAAGT CCAGGAGTGT 1560
 CCTGGAGAAG AAGAGGCGGC TCCTGGAGAA GCTGCAGGAG CAGCTGGCGC AGCTGAGTGT 1620
 GCAGGCCACG GACAAGGAGG AGAACAAGCA GGTGGCCCTG GGCACGTCCA AGCTCAACTA 1680
 CCTGGACCCC AGGATCAGCA TTGCCTGGTG CAAGCGGTTT AGGGTGCCAG TGGAGAAGAT 1740
 CTACAGCAAA ACACAGCGGG AGAGGTTTCG CTGGGCTCTC GCCATGGCAG GAGAAGACTT 1800
 TGAATTCTAA CGACGAGCCG TGTTGAAACT TCTTTTGTAT GTGTGTGTGT TTTTTTCACT 1860
 ATTAAAGCAG TACTGGGGAA TTTTGTACAA TAAAAA AAAA AAAA 1917

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 601 AMINO ACIDS
 (B) TYPE: AMINO ACID
 (C) STRANDEDNESS:
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Arg	Val	Val	Arg	Leu	Leu	Arg	Leu	Arg	Ala	Ala	Leu	Thr	Leu
			5						10					15
Leu	Gly	Glu	Val	Pro	Arg	Arg	Pro	Ala	Ser	Arg	Gly	Val	Pro	Gly
			20						25					30
Ser	Arg	Arg	Thr	Gln	Lys	Gly	Ser	Gly	Ala	Arg	Trp	Glu	Lys	Glu
			35						40					45
Lys	His	Glu	Asp	Gly	Val	Lys	Trp	Arg	Gln	Leu	Gly	His	Lys	Gly
			50						55					60
Pro	Tyr	Pag	Ala	Pro	Pro	Tyr	Glu	Pro	Leu	Pro	Asp	Gly	Val	Arg
			65						70					75
Phe	Phe	Tyr	Glu	Gly	Arg	Pro	Val	Arg	Leu	Ser	Val	Pro	Ala	Glu
			80						85					90
Glu	Val	Pro	Thr	Phe	Tyr	Gly	Arg	Met	Leu	Asp	His	Glu	Tyr	Thr
			95						100					105

Thr Lys Glu Val Phe Arg Lys Asn Phe Phe Asn Asp Trp Arg Lys	110	115	120
Glu Met Ala Val Glu Glu Arg Glu Val Ile Lys Ser Leu Asp Lys	125	130	135
Cys Asp Phe Thr Glu Ile His Arg Tyr Phe Val Asp Lys Ala Ala	140	145	150
Ala Arg Lys Val Leu Ser Arg Glu Glu Lys Gln Lys Leu Lys Glu	155	160	165
Glu Ala Glu Lys Leu Gln Gln Glu Phe Gly Tyr Cys Ile Leu Asp	170	175	180
Gly His Gln Glu Lys Ile Gly Asn Phe Lys Ile Glu Pro Pro Gly	185	190	195
Leu Phe Arg Gly Arg Gly Asp His Pro Lys Met Gly Met Leu Lys	200	205	210
Arg Arg Ile Thr Pro Glu Asp Val Val Ile Asn Cys Ser Arg Asp	215	220	225
Ser Lys Ile Pro Glu Pro Pro Ala Gly His Gln Trp Lys Glu Val	230	235	240
Arg Ser Asp Asn Thr Val Thr Trp Leu Ala Ala Trp Thr Glu Ser	245	250	255
Val Gln Asn Ser Ile Lys Tyr Ile Met Leu Asn Pro Cys Ser Lys	260	265	270
Leu Lys Gly Glu Thr Ala Trp Gln Lys Phe Glu Thr Ala Arg Arg	275	280	285
Leu Arg Gly Phe Val Asp Glu Ile Arg Ser Gln Tyr Arg Ala Asp	290	295	300
Trp Lys Ser Arg Glu Met Lys Thr Arg Gln Arg Ala Val Ala Leu	305	310	315
Tyr Phe Ile Asp Lys Leu Ala Leu Arg Ala Gly Asn Glu Lys Glu	320	325	330
Asp Gly Glu Ala Ala Asp Thr Val Gly Cys Cys Ser Leu Arg Val	335	340	345
Glu His Val Gln Leu His Pro Glu Ala Asp Gly Cys Gln His Val	350	355	360
Val Glu Phe Asp Phe Leu Gly Lys Asp Cys Ile Arg Tyr Tyr Asn			

	365		370		375
Arg Val Pro Val Glu Lys Pro Val Tyr		Lys Asn Leu Gln Leu Phe			
	380		385		390
Met Glu Asn Lys Asp Pro Arg Asp Asp		Leu Phe Asp Arg Leu Thr			
	395		400		405
Thr Thr Ser Leu Asn Lys His Leu Gln		Glu Leu Met Asp Gly Leu			
	410		415		420
Thr Ala Lys Val Phe Arg Thr Tyr Asn		Ala Ser Ile Thr Leu Gln			
	425		430		435
Glu Gln Leu Arg Ala Leu Thr Arg Ala		Glu Asp Ser Ile Ala Ala			
	440		445		450
Lys Ile Leu Ser Tyr Asn Arg Ala Asn		Arg Val Val Ala Ile Leu			
	455		460		465
Cys Asn His Gln Arg Ala Thr Pro Ser		Thr Phe Glu Lys Ser Met			
	470		475		480
Gln Aln Leu Gln Thr Lys Ile Gln Ala		Lys Lys Glu Gln Val Ala			
	485		490		495
Glu Ala Arg Ala Glu Leu Arg Arg Ala		Arg Ala Glu His Lys Ala			
	500		505		510
Gln Gly Asp Gly Lys Ser Arg Ser Val		Leu Glu Lys Lys Arg Arg			
	515		520		525
Leu Leu Glu Lys Leu Gln Glu Gln Leu		Ala Gln Leu Ser Val Gln			
	530		535		540
Ala Thr Asp Lys Glu Glu Asn Lys Gln		Val Ala Leu Gly Thr Ser			
	545		550		555
Lys Leu Asn Tyr Leu Asp Pro Arg Ile		Ser Ile Ala Trp Cys Lys			
	560		565		570
Arg Phe Arg Val Pro Val Glu Lys Ile		Tyr Ser Lys Thr Gln Arg			
	575		580		585
Glu Arg Phe Ala Trp Ala Leu Ala Met		Ala Gly Glu Asp Phe Glu			
	590		595		600
Phe					

WHAT IS CLAIMED IS:

1. An isolated polynucleotide selected from the group consisting of:
 - (a) a polynucleotide encoding an hTopI- α polypeptide having the deduced amino acid sequence of Figure 1 or a fragment, analog or derivative of said polypeptide;
 - (b) a polynucleotide encoding an hTopI- α polypeptide having the amino acid sequence encoded by the cDNA contained in ATCC Deposit No. 75714 or a fragment, analog or derivative of said polypeptide.
2. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.
3. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.
4. The polynucleotide of Claim 1 wherein the polynucleotide is genomic DNA.
5. The polynucleotide of Claim 2 wherein said polynucleotide encodes hTopI- α having the deduced amino acid sequence of Figure 1.
6. The polynucleotide of Claim 2 wherein said polynucleotide encodes an hTopI- α polypeptide encoded by the cDNA of ATCC Deposit No. 75714.
7. The polynucleotide of Claim 1 having the coding sequence of hTopI- α as shown in Figure 1.
8. The polynucleotide of Claim 2 having the coding sequence of hTopI- α deposited as ATCC Deposit No. 75714.
9. A vector containing the DNA of Claim 2.
10. A host cell genetically engineered with the vector of Claim 9.
11. A process for producing a polypeptide comprising: expressing from the host cell of Claim 10 the polypeptide encoded by said DNA.

12. A process for producing cells capable of expressing a polypeptide comprising genetically engineering cells with the vector of Claim 9.
13. An isolated DNA hybridizable to the DNA of Claim 2 and encoding a polypeptide having hTopI- α activity.
14. A polypeptide selected from the group consisting of (i) a hTopI- α polypeptide having the deduced amino acid sequence of Figure 1 and fragments, analogs and derivatives thereof and (ii) a hTopI- α polypeptide encoded by the cDNA of ATCC Deposit No. 75714 and fragments, analogs and derivatives of said polypeptide.
15. The polypeptide of Claim 14 wherein the polypeptide is hTopI- α having the deduced amino acid sequence of Figure 1.
16. An antibody against the polypeptide of claim 14.
17. An antagonist/inhibitor against the polypeptide of claim 14.
18. A method for the treatment of a patient having need to inhibit hTopI- α comprising: administering to the patient a therapeutically effective amount of the antagonist/inhibitor of Claim 17.
19. A pharmaceutical composition comprising the polypeptide of Claim 14 and a pharmaceutically acceptable carrier.
20. A method for determining an antagonist/inhibitor effective for preventing the action of hTopI- α comprising:
 - combining DNA, hTopI- α and a potential antagonist/inhibitor;
 - incubating the combination under effective conditions to allow for hTopI- α to act on the DNA; and
 - determining if the antagonist/inhibitor effectively blocks the action of hTopI- α .

```
569 CKFRVPVEKIYKTKQRRERFAWALAMAGEDFEF 601
    ||::||:|||||.|||||:|||||:..|||||:|
733 CKKWGVPIEKIYNKTKQREKFAWAIDMADEDYEF 765
```

0 . 20 40 60 80 100 120 140 160 180 200 220

GCAGATCGCGTGGTGGCGCTGCTCGCGCTCCGGCGCGCTCTGACGCTGCTCGGGGAGGT
CGTCTACGCGCACCGCGACGCGAGCGCGCGCGAGACTGCGACGAGCCCCCTCCA
M R V V R L L R L R A A L T L L G E V
CCCCCGCGCGCGCTCCCGGGGTGTCCTCCGGGCTCGCGCAGGACGAGGCGAGTGG
GGGGCGCGCGCGAGGCGCGCGCGCGCGCGCGCTCGCTTCCCGTCACC
P R R P A S R G V P G S R R T Q K G S G
AGCCAGGTGGGAGAGGAGAACGACGAGCGGGTGAAGTGAGACAGCTGGAGCACAA
TCGGTCCACCCCTCTTCCTCTTCGTGCTTCTGCCCCACTTCACCTCTGTCGACCTCGTGT
A R W E K E K H E D G V K W R Q L E H K
MATCH WITH FIG. 1B

FIG. 1B

MATCH WITH FIG. 1A

GGGCCCGTACTTCGCACCCCATACGAGCCCTTCCCGACGGAGTGCCTTCTCTATGA
 ---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
 CCGGGCATGAAGCGTGGGGTATGCTCGGGAAGGCTGCCTCAGCAAGAAGATACT
 G P Y F A P P Y E P L P D G V R F F Y E
 240 260 280

AGGAAGCCCTGTGAGATTGAGCGTGCCACGGAGGAGGTCCCACCTTTTATGGGAGGAT
 ---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
 TCCTTCGGGACACTCTAACTCGCACGGTGGCTCCCTCCAGGGTGAAATAACCCCTCCTA
 G R P V R L S V P A E E V P T F Y G R M
 300 320 340

GTTAGATCATGAATACACAACAAGGAGGTTTCCCGGAAGAACTTCTTCAATGACTGGCG
 ---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
 CAATCTAGTACTTATGTGTGTTTCCTCCAAAGGCCCTTCTTGAAGAAGTTACTGACCGC
 L D H E Y T T K E V F R K N F F N D W R
 360 380 400

AAAGGAAATGGCGGTGGAAGAGAGGGAAGTCATCAAGAGCCCTGGACAAGTGTGACTTCAC
 ---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
 TTTCCCTTACCGCCACCTTCTCTCCCTTCAGT'AGTTCTCGGACCTGTTCACTGAAGTG
 K E M A V E E R E V I K S L D K C D F T
 420 440 460

MATCH WITH FIG. 1C

2 / 1 0

SUBSTITUTE SHEET (RULE 26)

FIG. 1C

MATCH WITH FIG. 1B

GGAGATCCACAGATACTTTGTGGACAAGGCCGAGCCCGGAAGTCCCTGACGAGGAGGA
 ---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
 CCTCTAGGTGCTATGAACACCTGTTCCGGCGTGGGCCCTTTCAGGACTCGTCCCTCCT
 E I H R Y F V D K A A R K V L S R E E
 480 500 520

GAAGCAGAAGCTAAAGAAGAGGCAGAAAACCTTCAGCAAGAGTTCGGCTACTGTATTT
 ---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
 CTTCGCTCTCGATTTCTCTCCGCTTTTGAAGTCGTTCTCAAGCCGATGACATAAAA
 K Q K L K E E A E K L Q Q E F G Y C I L
 540 560 580

AGATGGTCACCAAGAAAAATAGGCAACTTCAAGATTGAGCCGCCCTGGCTTGTCCGTGG
 ---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
 TCTACCAAGTGGTCTTTTATCCGTTGAAGTTCTAACTCGGCGGACCGAACAAGGCACC
 D G H Q E K I G N F K I E P P G L F R G
 600 620 640

CCGTGGCGACCATCCCAAGATGGGATGCTGAAGAGAAGGATCACGCCAGAGGATGTGGT
 ---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
 GGCACCGCTGTAGGTTCTACCCCTACGACTTCTCTCCTAGTGGGTCTCCTACACCA
 R G D H P K M G M L K R R I T P E D V V
 660 680 700

MATCH WITH FIG. 1D

4 / 1 0

— 6 —

1000

1060

1120

1180

MATCH WITH FIG. 1F

MATCH WITH FIG. 1E FIG. 1F

```

      . . . . .
GGAGAAGCCGGTGTAACAAGAACTTACAGCTCTTTATGGAGAACAGGACCCCGGACGA
-----+-----+-----+-----+-----+-----+-----+-----+
CCTCTTCGGCCACATGTTCTTGAATGTCGAGAAATACCTCTTGTTCCTGGGGGCCCTGCT
E K P V Y K N L Q L F M E N K D P R D D
1200 . 1220 1240

      . . . . .
CCTCTTCGACAGGCTGACCACGACGACCCTGAACAAGCACCTCCAGGAGCTGATGGACGG
-----+-----+-----+-----+-----+-----+-----+
GGAGAAGCTGTCCGACTGGTGGTGGTGGACTTGTTCGTGGAGGTCTCGACTACCTGCC
L F D R L T T T S L N K H L Q E L M D G
1260 1280 1300

      . . . . .
GCTGACGGCCAAGGTGTTCCGGACCCTACAACGCCCTCCATCACTCTGCAGGACGCTGCC
-----+-----+-----+-----+-----+-----+-----+
CGACTGCCCGGTTCCACAAGGCCCTGGATGTTGCCGGAGGTAGTGAGACGTCCTCGTCGACGC
L T A K V F R T Y N A S I T L Q E Q L R
1320 1340 1360

      . . . . .
GGCCCTGACGGCGCGGAGGACAGCATAGCAGCTAAGATCTTATCCTACAACCGAGCCAA
-----+-----+-----+-----+-----+-----+-----+
CCGGGACTGCCGCGGGCTCCTGTGCTATCGTCGATCTAGAATAGGATGTTGGCTCGGTT
A L T R A E D S I A A K I L S Y N R A N
1380 1400 1420

```

MATCH WITH FIG. 1G

4-6-61

MATCH WITH FIG. 1F

CCGAGTCGTGGCCATTCTCTGCCAACCATCAGCGAGCAACCCAGTACGTTCCGAGAAGTC
 GGCTCAGCACCGGTAAGAGACGTTGGTAGTCGCTCGTTGGGGTCATGCAAGCTCTTCAG
 R V V A I L C N H Q R A T P S T F E K S
 1440 1460 1480

GATGCAGATCTCCAGACGAGATCCAGGCCAAGAGGAGCAGGTGCTGAGGCCAGGGC
 CTACGCTTTAGAGGTCGTCTTAGGTCCGTTTCTTCCTCGTCCACGACTCCGGTCCCG
 M Q N L Q T K I Q A K K E Q V A E A R A
 1500 1520 1540

AGAGCTGACGAGCGGCTGAGCACAAAGCCCAAGGGATGGCAAGTCCAGGAGTG
-----+-----+-----+-----+-----+-----+-----
TCTCGACTCCTCCGCTCCGACTCGTGTTTCGGGTCCCCCTACCGTTCAGGTCCTCACA
E L R R A R A E H K A Q G D G K S R S V'
1560 1580 1600

CCTGGAGAAGAGCGGCTCCTGGAGAAGCTGCAGGACGCTGGCCGAGCTGAGTG
-----+-----+-----+-----+-----+-----+-----
GGACCTCTTCTTCGCCGAGGACCCTCTTCGACGTCTCTCGTCCGCCGCTCGACTCACA
L'E K K R R L L E K L Q E Q L A Q L S V

MATCH WITH FIG. 1H

MATCH WITH FIG. 1G

1620 . 1640 . 1660 .
GCAGGCCACGGACAAGGAGGAGAAACAAGCAGGTGGCCCTGGGCACGTCCTCAAGCTCAACTA
CGTCCGGTGCCCTGTTCCTCCTCTTGTTCGTCCACCGGACCCGTCAGGTCGAGTTGAT
Q A T D K E E N K Q V A L G T S K L N Y
1680 . 1700 . 1720 .
CCTGACCCAGGATCAGCATTCCTGCTGCAAGCGGTTCAGGGTGCCAGTGCGAGAAGAT
GGACCTGGGGTCCCTAGTCGTAACGGACACACGTTCCGCCAAGTCCACGGTCACCTCTTCTA
L D P R I S I A W C K R F R V P V E K I
1740 . 1760 . 1780 .
CTACAGCAAAACACAGCGGAGAGGTTCCGCTTGGCTCTCGCCATGGCAGGAGAAGACTT
GATGTCGTTTGTGTCGCCCTCTCCAAGCGACCCGAGAGCGGTACCGTCTCTTCTGTAA
Y S K T Q R E R F A W A L A M A G E D F
1800 . 1820 . 1840 .
TGAATTCTAACGACGACCGGTGTGAAACTTCTTTGTATGTGTGTGTTTCTTCTCACT
ACTTAAGAT^{*}TGCTGCTCGGCACAACTTTGAAGAAACATACACACACAAAAAAGTGA
E_F 1860 . 1880 . 1900 .
ATTAAGCAGTACTGGGAATTGTGTACAAATAAAAAAATAAAAAAATAAAAAA
TAATTTCGTCATGACCCCTTAAACATGTTATTTTCTTTTCTTTTCTTTTCTTTTCTTTT

FIG. 2A

[illegible]

MATCH WITH FIG. 2B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/05701

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C12N 9/90, 15/61; A61K 37/58

US CL : 435/233; 536/23.2; 424/94.5

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/233; 536/23.2; 424/94.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Computer Search - CA and APS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US, A, 5,070,192 (EARNSHAW ET. AL.) 03 December 1991, see entire document.	1-15 and 19
A	Cancer Res., Volume 52, issued 01 February 1992, K. R. Madden et. al., "Overexpression Of Human Topoisomerase I In Baby Hamster Kidney Cells: Hypersensitivity Of Clonal Isolates To Camptothecin", pages 525-532.	1-15 and 19
A	Chemical Abstracts, Volume 121, issued 29 August 1994, D. S. Samuels et. al., "The Predominant Form Of Mammalian DNA Topoisomerase I In Vivo Has A Molecular Mass Of 100 kDA", page 428, column 1, The Abstract 102667g, Mol. Biol. Rep. 1994, 19(2), 99-103.	1-15 and 19

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

18 OCTOBER 1994

Date of mailing of the international search report

DEC 02 1994

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Authorized officer

Charles Patterson

Facsimile No. (703) 305-3230

Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/05701

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Proc. Natl. Acad. Sci. USA, Volume 85, issued April 1988, P. D'Arpa et. al., "cDNA cloning of human DNA topoisomerase I: Catalytic activity of a 67.7 kDa carboxyl-terminal fragment", pages 2543-2547.	1-15 and 19

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/05701

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
1-15 and 19
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/05701

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

- I. Claims 1-13, drawn to a polynucleotide, a vector containing to polynucleotide, a host cell containing the vector and a method of use.
- II. Claims 14-15 and 19, drawn to a polypeptide.
- III. Claim 16, drawn to an antibody.
- IV. Claims 17, 18 and 20, drawn to an antagonist/inhibitor, a method of use and a method for assaying for the antagonist/inhibitor.

The inventions do not meet the requirements of unity of invention because there is no technical relationship among the inventions involving one or more of the same or corresponding special technical features, as defined in PCT Rule 13.2, so as to form a single inventive concept. The polynucleotide, polypeptide, antibody and antagonist of Groups I-IV, respectively, are separate and distinct products.